УДК 519.854

С. С. ЛУГОВОЙ, студент НТУ «ХПИ»,
Ю. С. ШАХНОВСКИЙ, канд. техн. наук, доцент НТУ «ХПИ»

АНАЛИЗ УСЛОВИЙ ДОПУСТИМОСТИ ПЕРЕСТАНОВКИ ПРАВИЛ МЕЖДУ ЦЕПОЧКАМИ В СИСТЕМЕ ФИЛЬТРАЦИИ ПАКЕТОВ

У статті розглядається оптимізація систем фільтрації Internet-шлюзів. Запропоновані додаткові резерви оптимізації за рахунок перестановки правил між ланцюжками. Сформульовано умови припустимості перестановки правил та правило розрахунку критерію якості для нової перестановки.

В статье рассматривается оптимизация систем фильтрации Internet-шлюзов. Предложены дополнительные резервы оптимизации за счет перестановки правил между цепочками. Сформулированы условия допустимости перестановки правил и правило расчета критерия качества для новой перестановки.

In article optimization of systems of a filtration of Internet-gateways is considered. Additional reserves of optimization are offered due to rearrangement of rules between chains. Conditions of an admissibility of rearrangement of rules and a rule of calculation of criterion of quality for new rearrangement are formulated.

Введение. В сетевых шлюзах для управления прохождением пакетов через шлюз используются системы фильтрации. Такие системы состоят из нескольких упорядоченных последовательностей (цепочек) правил фильтрации. Существует три типа правил. Пакет, попавший в систему, поступает на проверку главной цепочкой правил фильтрации. Проверка правилами осуществляется в порядке их следования в цепочке. При нахождении соответствия данного пакета некоторому правилу типа ***а*** пакет покидает систему фильтрации (будет переслан далее по маршруту либо уничтожен). При соответствии данного пакета правилу типа ***б*** дальнейшая судьба данного пакета будет определяться правилами цепочки, на которую передается управление. Правило типа ***в*** определяет передачу пакета на проверку в предыдущую (вызывающую) цепочку. Цепочки соединены между собой правилами типа ***б*** и ***в***. Если построить орграф, вершинами которого будут цепочки, а дугами – передачи управления правилами типа ***б***, он будет ациклическим.

Чем меньшее количество правил проверяется для принятия решения о судьбе пакета, тем меньше затраты вычислительного времени ЭВМ. Допустимо переупорядочивание правил с сохранением результата работы системы. В [1] сформулированы условия, когда переупорядочивание правил внутри одной цепочки разрешено. За счет этого можно добиться уменьшения суммарного количества проверок всех пакетов системой фильтрации.

В [1, 2] рассмотрена задача оптимизации затрат времени при переупорядочивании правил в пределах каждой цепочки системы фильтрации

в отдельности, а также разработаны точные и эвристические алгоритмы решения поставленной задачи. Однако в этих работах не учитывалась возможность перестановки правил из одной цепочки в другую, что может дать дополнительные резервы для оптимизации.

Существующие системы фильтрации накапливают информацию о числе пакетов, соответствующих каждому правилу. Разумно предположить, что на следующем интервале времени относительные частоты пакетов изменяться не сильно. Время расчета по каждому правилу зависит от применяемых для расчета программ, однако в большинстве случаем это время можно считать единичным.

Постановка задачи. Известны: а) относительная частота встречи пакетов, соответствующих каждому из правил; б) список цепочек, составляющих систему фильтрации, а также списки правил, входящих в каждую из цепочек. Необходимо найти такую последовательность правил в цепочках, при которой среднее время прохождения пакета через систему фильтров минимально.

Критерии допустимости. Для решения поставленной задачи необходимо определить критерии допустимости перестановки правил из одной цепочки в другую. Введем обозначения, характеризующие каждое из правил, входящих в систему фильтрации.

Пронумеруем все цепочки правил и отдельные правила внутри цепочек арабскими цифрами, например, правило (i, j) – j-е правило i-й цепочки. При перестановке правил будем эти номера сохранять даже в случае изменения порядка следования правил. Т.е. номера правил полностью определяются их положением в начальной перестановке. Синтаксис каждого правила определяет множество пакетов, выделяемых данным правилом из универсума (всего теоретически возможного множества пакетов). Обозначим множество, задаваемое синтаксисом правила (i, j) как $P_{i,j}$. Но на вход правила поступает не универсум, а множество, ограниченное предшествующими правилами системы. Обозначим это множество как $U_{i,j}$. Множество пакетов, которые реально будут обработаны этим правилом, обозначим как $O_{i,j} = U_{i,j} \cap P_{i,j}$. Так же необходимо ввести множество пакетов, которые покинут систему после обработки (i, j)-м правилом $L_{i,j}$. Для правил типа ***а*** данное множество совпадает с $O_{i,j}$, для правил типа ***в*** оно является пустым множеством, а для правила типа ***б*** данное множество можно рассчитать по формуле:

$$L_{i,j} = O_{i,j} \cap \bigcup_{m \in M} L_{d,m} ,$$

где d – номер цепочки, на которую ссылается правило (i, j);

M – множество правил цепочки d.

Соотношение между введенными множествами показано на рис. 1.

Определим теперь условия допустимости перестановки правил из одной цепочки в другую. Будем рассматривать перестановку только самого первого правила дочерней цепочки и правила, непосредственно стоящего перед правилом перехода в дочернюю цепочку. Т.к. из [1] следует, что допускается перестановка правила из дочерней цепочки в родительскую только в том случае, если при перестановке в пределах дочерней цепочки оно может быть поставлено на первую позицию, а так же допускается перестановка правила из родительской цепочки в дочернюю только в том случае, если при перестановке в пределах родительской цепочки, оно может быть поставлено непосредственно перед правилом ветвления. Так же необходимо отметить, что допускается перестановка из одной цепочки в другую только правил типа ***а*** и ***б***.

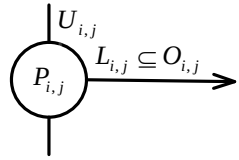

Рис. 1

Поскольку любая цепочка задает множество пакетов, покидающих систему, являющееся объединением множеств L, задаваемых входящими в нее правилами, то условия перестановки для правил типов ***а*** и ***б*** одинаковы.

Для наглядности на рисунках, поясняющих принципы перестановки правил, будем отображать состояние системы в начальный момент времени и в момент после перестановки правила.

Рассмотрим несколько основных структур связей между цепочками и условия допустимости перестановки правил из одной цепочки в другую для данного типа структуры. Несложно показать, что любые другие возможные структуры связей являются либо частным случаем основной структуры, либо комбинацией из нескольких основных структур.

Проанализируем эти структуры. Существует два варианта.

А) Одна цепочка является дочерней по отношению к нескольким другим (рис. 2, 3).

Как показано на рис. 2, правило дочерней цепочки d вставляем во все родительские при совместном выполнении условий:

$$U_{im,jm} \cap P_{d,1} \subseteq P_{im,jm}, m = 1..n.$$

После этого удаляем вставленное правило из родительских цепочек, для которых выполняется условие:

$$O_{im,jm} \cap P_{d,1} = \varnothing,$$

причем для проверки данного условия берем $O_{im,\,jm}$, рассчитанное до осуществления перестановки.

Обратная перестановка правила родительской цепочки в дочернюю показана на рис. 3.

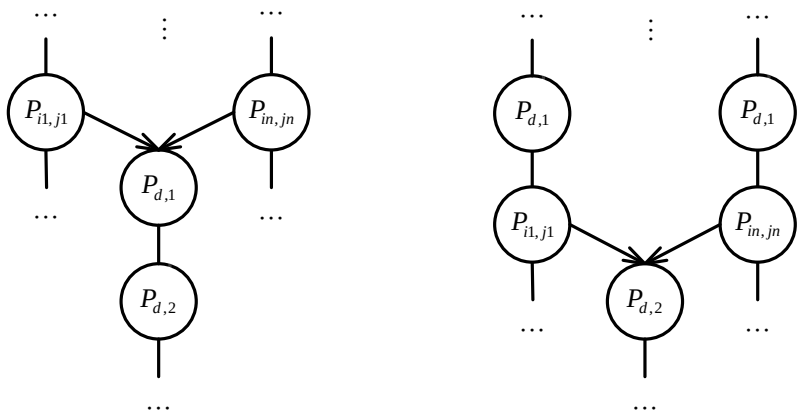

Рис. 2

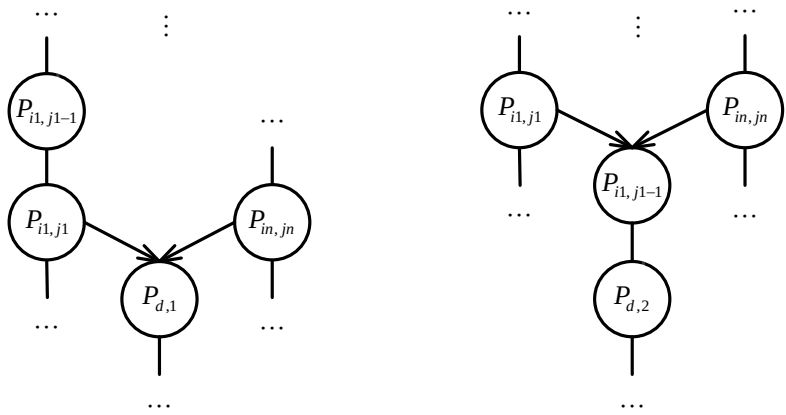

Рис. 3

В данном случае перестановка правила $(i1,\, j1-1)$ из родительской цепочки $i1$ в дочернюю цепочку d допустима при выполнении условий:

$$\begin{cases} O_{im,\,jm} \cap P_{i1,\,j1-1} = \varnothing, m = 2..n, \\ L_{i1,\,j1-1} \subseteq P_{i1,\,j1}. \end{cases}$$

Б) Несколько правил одной и той же родительской цепочки определяют переход в одну дочернюю цепочку (рис. 4, 5).

По рис. 4 проводим вставку правила $(d,1)$ перед всеми правилами родительской цепочки, определяющими переход в дочернюю цепочку d, для которых выполняется условие:

$$U_{i,jm} \cap P_{d,1} \subseteq P_{i,jm}, m = 1..n.$$

Далее в цепочке i оставляем только то правило $(d,1)$, которое было вставлено перед правилом с минимальным номером jm.

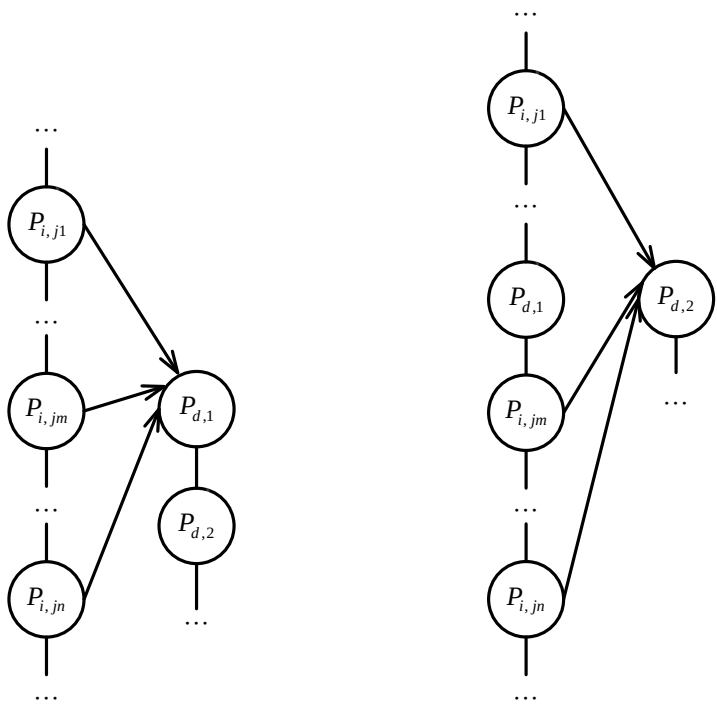

Рис. 4

При обратной перестановке правила из родительской цепочки в дочернюю, необходимо проверить условие допустимости перестановки:

$$\begin{cases} O_{i,jk} \cap P_{i,jm-1} = \varnothing, k = 1..m-1, \\ L_{i,jm-1} \subseteq P_{i,jm}. \end{cases}$$

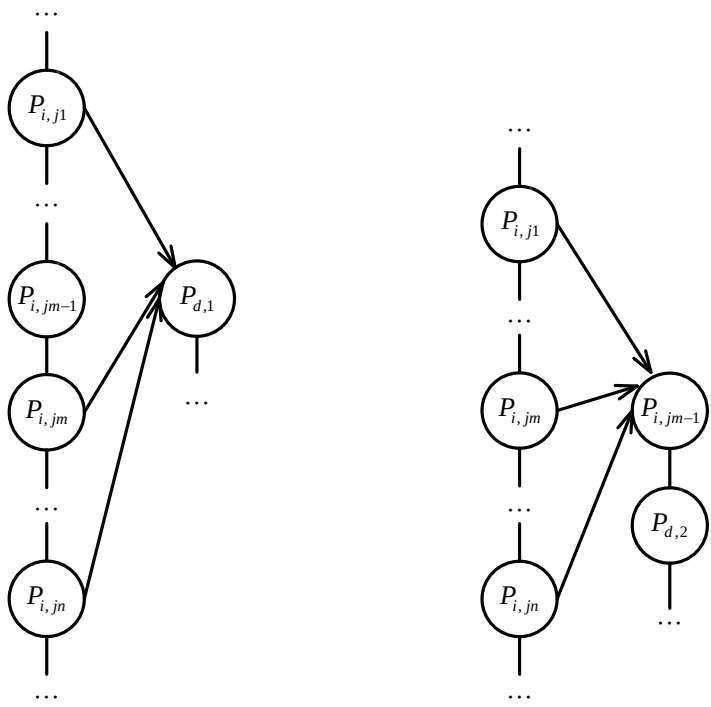

Рис. 5

Критерий качества. В процессе решения задачи оптимизации необходимо уметь сравнивать различные решения. Поскольку в задаче оптимизируются затраты времени, критерием качества после изменения порядка будет суммарное время проверки всех пакетов системой. Необходим алгоритм расчета этой величины. Если нам известно количество пакетов соответствующих множеству $U_{i,j}$ для любого правила (i, j), то можно рассчитать критерий по формуле:

$$K = \sum_{i=1}^{n} \sum_{j=1}^{m_i} k_{i,j} ,$$

где K – искомый критерий качества;

n – количество цепочек в системе;

m_i – количество правил в цепочке i;

$k_{i,j}$ – количество пакетов, соответствующее множеству $U_{i,j}$.

Попытаемся определить точное значение критерия качества некоторой простой системы цепочек, представленной рис. 6. Обозначим количество

пакетов, прошедших по некоторому направлению строчными латинскими буквами $a..g$, как показано на рисунке. Пусть правило $(2,1)$ является правилом типа ***a***. Тогда количество пакетов, отвечающее правилу $(2,1)$, f состоит из двух слагаемых f_1 и f_2 – обработанных по ветвлению по $(1,1)$ и $(1,2)$ правилам соответственно:

$$f = f_1 + f_2 .$$

Аналогично для пакетов, не соответствующих правилу $(2,1)$:

$$g = g_1 + g_2 .$$

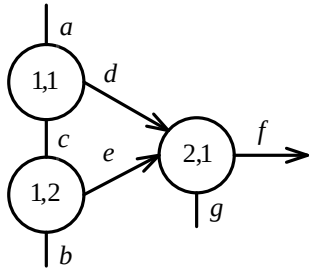

Рис. 6

Тогда критерий качества составит:

$$K = a + c + (d + e).$$

Причем a, d и e известны, а c – нет. Для нахождения неизвестного c запишем систему уравнений относительно $a..g$, которые следуют из приведенной системы правил:

$$\begin{cases} c = (a - d) + g_1, \\ b = (c - e) + g_2. \end{cases}$$

Параметры g_1 и g_2 неизвестны, но связаны равенством, добавим его в систему уравнений и получим:

$$\begin{cases} c = (a - d) + g_1, \\ b = (c - e) + g_2, \\ g = g_1 + g_2. \end{cases}$$

Несложными математическими преобразованиями можно показать, что данная система уравнений является линейно-зависимой, и нет возможности определить неизвестное c.

Это означает, что информации выдаваемой существующими системами фильтрации не достаточно для точного расчета сформулированного критерия качества перестановки правил. В рамках данной работы не предполагается вносить изменения в механизм сбора информации системой фильтрации. Вместо этого предлагается метод получения приблизительной оценки качества системы правил.

Причиной невозможности точного расчета критерия качества является отсутствие полной информации о соотношении количества пакетов из родительских цепочек, которые соответствуют множеству O правила из дочерней цепочки. Предлагается считать, что количество пакетов f_1, покидающее систему по некоторому правилу дочерней цепочки, пропорционально отношению количества пакетов обработанных правилом родительской цепочки, по которому происходит переход в данную дочернюю цепочку (d), к общему количеству пакетов попавших в цепочку $(d+e)$. Данное допущение вносит в оценку критерия качества погрешность. Для оценки этой погрешности был проведен эксперимент на 100 случайно сгенерированных системах цепочек. Средняя погрешность расчета по предложенному правилу составила 0.92 %.

Эксперименты, проведенные с оптимизацией внутри цепочки [2] дали выигрыш значительно больше этой величины. С другой стороны погрешность, вносимая изменением во времени частот пакетов, соответствующих правилам, даст погрешность большую, чем погрешность предложенного выше метода. Поэтому предлагаемый метод расчета критерия качества применим.

Выводы. Сформулированы критерии допустимости перестановки правил между цепочками. Предложен метод расчета критерия качества системы фильтрации, состоящей более чем из одной цепочки с использованием данных, собираемых существующими системами фильтрации. На основе этих исследований необходимо разработать алгоритмы оптимизации системы цепочек, которые должны дать больший выигрыш по сравнению с алгоритмами оптимизации внутри цепочки, предложенных в [1, 2].

Список литературы: ***1.*** *Шахновский Ю. С., Гончаров А. В.* Оптимизация системы фильтрации пакетов.– Вестник НТУ «ХПИ». Сб. научных трудов. Тем. вып. «Системный анализ, управление и информационные технологии».– Харьков: НТУ «ХПИ».– 2003.– № 6.– С. 53-56. ***2.*** *Шахновский Ю. С.* Оптимизация порядка последовательности условий – метод ветвей и границ. – Вестник НТУ «ХПИ». Сб. научных трудов. Тем. вып. «Системный анализ, управление и информационные технологии».– Харьков: НТУ «ХПИ».– 2003.– № 18.– С. 33-36.

Поступила в редколлегию 25.11.08